



## **Parsimonious Short-Term Load Forecasting for Optimal Operation Planning of Electrical Distribution Systems**

**López, Juan Camilo ; Rider, Marcos J. ; Wu, Qiuwei**

*Published in:*  
I E E E Transactions on Power Systems

*Link to article, DOI:*  
[10.1109/TPWRS.2018.2872388](https://doi.org/10.1109/TPWRS.2018.2872388)

*Publication date:*  
2018

*Document Version*  
Peer reviewed version

[Link back to DTU Orbit](#)

*Citation (APA):*  
López, J. C., Rider, M. J., & Wu, Q. (2018). Parsimonious Short-Term Load Forecasting for Optimal Operation Planning of Electrical Distribution Systems. I E E E Transactions on Power Systems, 34(2), 1427 - 1437.  
<https://doi.org/10.1109/TPWRS.2018.2872388>

---

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Parsimonious Short-Term Load Forecasting for Optimal Operation Planning of Electrical Distribution Systems

Juan Camilo López, Marcos J. Rider, *Senior Member, IEEE*, and Qiuwei Wu, *Senior Member, IEEE*

**Abstract**—The optimal operation planning (OOP) of electrical distribution systems (EDS) is very sensible to the quality of the short-term load forecasts. Assuming aggregated demands in EDS as univariate non-stationary seasonal time series, and based on historical measurements gathered by smart meters, this paper presents a parsimonious short-term load forecasting method to estimate the expected outcomes of future demands, and the standard deviations of forecast errors. The chosen short-term load forecasting method is an adaptation of the multiplicative autoregressive integrated moving average (ARIMA) models. Seasonal ARIMA models are parsimonious forecasting techniques because they require very few parameters and low computational resources to provide an adequate representation of stochastic time series. Two approaches are used in this paper to estimate the parameters that constitute the proposed multiplicative ARIMA model: a frequentist and a Bayesian approach. Advantages and disadvantages of both methods are compared by simulating a centralized self-healing scheme of a real EDS that uses the forecasts to deploy a robust restoration plan. Results shown that the proposed seasonal ARIMA model is a fast, precise, straightforward and adaptable load forecasting method, suitable for OOP of highly supervised EDS.

**Index Terms**—Electrical distribution systems, frequentist and Bayesian approach, optimal operation planning, seasonal ARIMA models, short-term load forecasting.

## I. INTRODUCTION

SHORT-TERM load forecasting is of great importance in the operation planning of bulk energy systems. Regulation bids, energy arbitrage, and market-clearing mechanisms are conducted on hourly bases, which puts a lot of pressure on forecasting techniques to provide accurate and fast estimations of future demands [1]. With the advent of smart meters and advanced distributed automation, electrical distribution system (EDS) operators are also becoming active users of short-term operation planning [2]. Based on the information gathered by meters, demand forecasts can be used as input data for the dynamic optimization of the EDS resources, a.k.a., optimized energy management systems. Moreover, in case of a fault, fast restoration methods can be deployed to minimize the total amount of expected unsupplied demand while the fault is being repaired, a.k.a., self-healing schemes [3].

J. C. López, and M. J. Rider are with the Department of Energy and Systems, UNICAMP – University of Campinas, Campinas, CEP 13083-852, Brazil. (e-mail: jclopeza@dsee.fee.unicamp.br; mjrrider@dsee.fee.unicamp.br).

Q. Wu is with the Center for Electric and Energy, Department of Electrical Engineering, Technical University of Denmark (DTU), Kgs. Lyngby 2800, Denmark, and Harvard China Project, School of Engineering and Applied Sciences (SEAS), Harvard University, 29 Oxford Street, Cambridge, 02138 MA, USA (e-mail: qw@elektro.dtu.dk, qiuwu@seas.harvard.edu).

Optimal operation planning (OOP) problems, such as self-healing schemes, require fast and precise short-term load forecasts in order to compute their decisions. In many cases, only the expected value of the future consumption is not enough. Other statistical moments, such as the standard deviation of forecast errors, are also important to make robust and risk-averse decisions. Such is the case for self-healing schemes, in which EDS operators would be more interested in a restoration plan that not only minimizes the expected amount of unsupplied demand, but also guarantees that the energized portion of the network is able to operate within its operational constraints during the time required to repair the fault. Thus, when implementing a self-healing scheme, the accuracy of the expected outcomes, and the standard deviations of the forecast errors, are significant to the robustness of the final restoration plan [4].

Since the beginnings of electrical engineering, numerous short-term load forecasting techniques have been proposed and extensively tested [5], [6]. Most techniques use historical information to formulate and train different estimators: multiple linear regression models [7], semi-parametric additive models [8], [9], exponential smoothing models [10], autoregressive moving average models [11], [12], artificial neural networks [13], [14], fuzzy regression models [15], [16], support vector machines [17], and gradient boosting [18], [19]; are among the most successful forecasting techniques proposed to this date.

In this context, the objective of this paper is not to propose a new short-term load forecasting technique in an already saturated research area. Our goal is, in fact, to adapt one of the aforementioned methods for formulating and solving short-term OOP problems in highly supervised EDS. To achieve this, the forecasting technique must meet the following four requirements: a) its computational complexity must be low in order to be deployed in on-line on-site applications; b) it should be accurate in terms of expected values and residuals; c) it should contain as few tuning parameters and forecasting variables as possible; and d) it should be able to automatically adapt itself to the ever-changing stochastic nature of EDS demands. A short-term load forecasting technique that satisfies all these four requirements is said to be a *parsimonious* method [20].

As shown in previous empirical studies, multiplicative autoregressive integrated moving average (ARIMA) models are simplistic and accurate analytical methods for short-term forecasting of aggregated demands [1]. All forecasting techniques are subject to error but, when dealing with univariate seasonal

stochastic time series, regression models are regarded as the best option to provide fast and reliable forecasts [21]. Assuming aggregated demands as non-stationary seasonal time series, several univariate multiplicative ARIMA models (one for each electrical measurement) can be used to dynamically estimate future outcomes of power demands in each load node. In this paper, a generalized parsimonious short-term load forecasting method based on seasonal ARIMA models and historical measurements gathered by smart meters is presented. Two approaches are used to estimate the parameters that constitute the proposed multiplicative ARIMA model: a frequentist and a Bayesian approach. The accuracy and efficiency of both estimation methods are tested by simulating a robust self-healing scheme in a real EDS that requires fast and precise short-term load forecasts to deploy its restoration plan. Results show that the proposed seasonal ARIMA model satisfies the four parsimonious requirements, and it is suitable for solving short-term OOP problems in supervised EDS.

The main contributions of this paper are as follows:

- A parsimonious short-term load forecasting method, based on seasonal ARIMA models and historical measurements gathered by smart meters, is adapted and presented as the first-step for solving short-term OOP problems in highly supervised and automated EDS.
- A frequentist and a Bayesian approaches for estimating the parameters of the seasonal ARIMA models are used and compared. Advantages and disadvantages of both methods are discussed by simulating a centralized self-healing scheme that uses the forecasted demands to execute a robust restoration plan.

The rest of the paper as organized as follows: Section II presents the application of forecasting techniques for the OOP of highly supervised EDS. Section III shows the process of model identification, preliminary estimation and forecasting, using the proposed seasonal ARIMA model. Section IV deals with the estimation of the model's parameters using the frequentist and Bayesian approaches. Model adequacy is discussed in Section V, and test and results are shown in Section VI, followed by conclusions.

## II. SHORT-TERM LOAD FORECASTING FOR THE OPTIMAL OPERATION PLANNING OF EDS

The optimal operation planning (OOP) of EDS requires fast, updated, and precise load forecasts to make informed decisions [6]. As shown in Fig. 1, short-term load forecasting is the very first step for deploying an OOP method in distribution systems with an advanced metering infrastructure (AMI) [22]. Based on current measurements (i.e., at time  $\tau$ ) and historical information, the short-term load forecasting method (the ARIMA model in Fig. 1) estimates the expected values and the standard deviations of future demands at each load node, i.e., from time  $\tau$  to  $\tau + l$ , where  $l$  is the time lag of the forecasts.

Electrical measurements, such as current and voltage magnitudes at distribution transformers, can be considered as non-stationary seasonal time series that depend on the aggregated demands of all users connected to each transformer. Thus, as shown in Fig. 1, these time series can be forecasted on-line

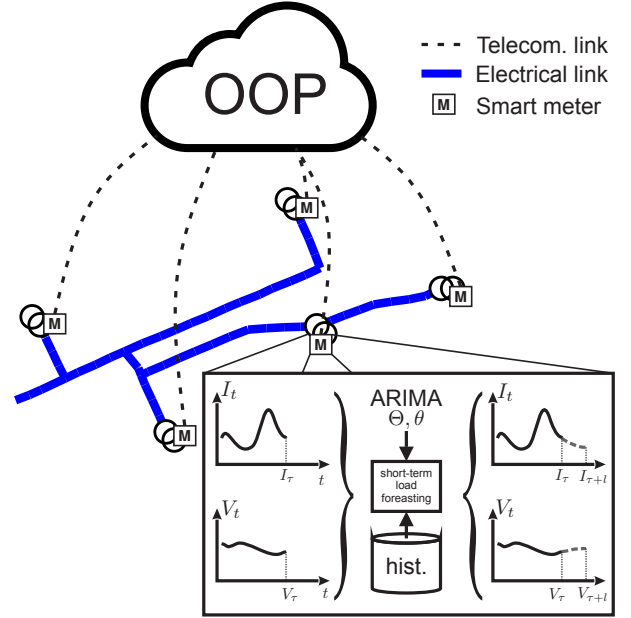


Figure 1. Short-term load forecasting for OOP of highly supervised EDS.

and on-site by taking advantage of the limited storage and processing capabilities of smart meters in the field.

Using a small resolution (e.g., 5 minutes per sample), the short-term memory of regression models can capture slow consumption patterns affected by exogenous variables, such as temperature and humidity [1]. Moreover, control actions that affect the collective consumption in a seasonal fashion (e.g., demand response at certain hours of the day, scheduled charging of EVs, etc.) can be also captured by univariate seasonal methods.

Finally, it is worth mentioning that univariate load forecasting methods should be used to estimate pure aggregated consumption, i.e., their are neither suitable for forecasting individual demands (e.g., households) nor renewable generation resources, because these stochastic processes are highly influenced by rapid exogenous random events, such as micro-weather conditions or individual human behavior, that cannot be fully captured by univariate linear regression models [2].

### A. OOP problem: Self-healing Scheme

In order to evaluate the efficiency of the proposed parsimonious load forecasting technique, a real-size EDS with an AMI will be used to simulate a centralized self-healing scheme. In case of a permanent fault, the function of the self-healing scheme is to automatically generate a set of control actions that minimize the impact of the outage while the fault is being repaired. Based on the short-term load forecasts and the location of the fault, the self-healing scheme is represented as an OOP problem whose solution aims at minimizing the expected unsupplied demand after the fault [23], [24].

The quality of the self-healing scheme is highly influenced by the speed and precision of the load forecasts. The set of restorative actions must be deployed shortly after the fault has been identified in order to be effective. Thus, forecasts must be updated and available at any time. Moreover, poor forecasting can result in ill-conceived restoration plans. For example, too much load could be transferred to heavily loaded feeders or, on

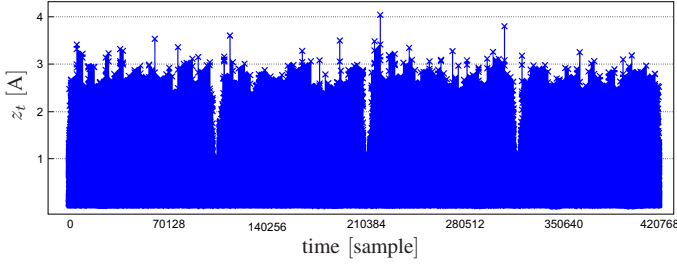


Figure 2. Time series  $z_t$ : four years of single-phase current magnitudes recorded from the primary side of a distribution transformer.

the contrary, feeders with enough capacity could be ignored by the self-healing scheme. Thus, in this case, not only the average values of the load forecasts are important, but also the standard deviation of the forecast errors are significant to the robustness of the restoration.

In this paper, the optimization model in [25] is used to represent and deploy the restoration plan of the centralized self-healing scheme in unbalanced three-phase EDS. The model in [25] considers the operation of remotely-controlled switches for transferring de-energized sections of the system to other feeders and for isolating the fault. Also, if available, dispatchable distributed generation (DG) units can be rescheduled to improve the impact of the restoration plan. The optimization model is shown in the Appendix. Details regarding the optimization technique used to solve the restoration plan will be discussed in Section VI.

### III. SEASONAL ARIMA MODEL

In this section, the use of a seasonal ARIMA model as a parsimonious short-term load forecasting technique is justified by analyzing the estimated autocorrelation function of a typical time series, associated to the magnitude of an aggregated single-phase demand current. Furthermore, the identification of the model, the initial estimation of its parameters, and the generation of forecasts will be discussed and validated through empirical analysis.

#### A. Model Identification

The first step of building any stochastic model is to analyze the attributes of the time series that will be forecasted. The time series in Fig. 2 represents the single-phase current magnitudes recorded by a smart meter at the primary side of a distribution transformer. Four years of data with a resolution of 5 minutes per sample are shown in Fig. 2, which constitutes a total amount of 420,768 measurements. The data have been generated by aggregating several demands from a residential area whose individual load profiles have been randomly generated using the LoadProfileGenerator software in [26], [27].

Time series  $z_t$  in Fig. 2 is a non-stationary stochastic process since its mean level and variance may have different values for different, sufficiently large, time intervals. However, there is a notorious interdependence between adjacent observations of the series which indicates that, in spite of its randomness, aggregated demands have a tendency of following similar patterns that derive from previous outcomes, i.e., there are temporal correlations among data samples. Analyzing these correlations is the first step for model identification. Fig. 3

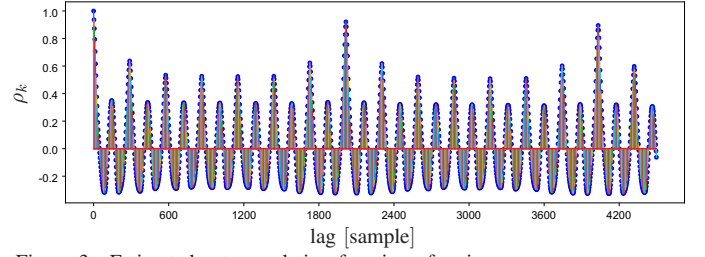


Figure 3. Estimated autocorrelation function of series  $z_t$ .

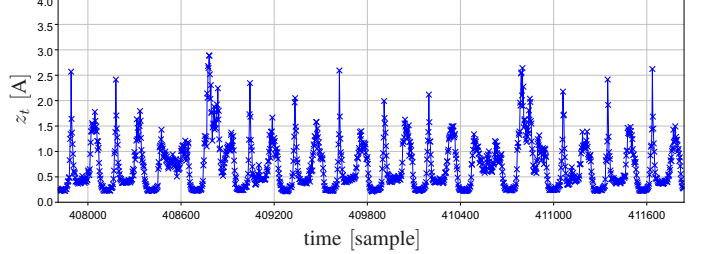


Figure 4. Detailed section of time series  $z_t$  that shows its seasonality.

shows the estimated autocorrelation function of series  $z_t$  for time lags between 1 to 4,500 samples (i.e., up to two weeks).

The estimated autocorrelation function in Fig. 3 seems to oscillate without visible reduction, with two notorious spikes at lags 2,016 and 4,032, i.e., at exactly one and two weeks, which suggests a periodic time series with period  $s = 2,016$ . Fig. 4 shows a detailed section of time series  $z_t$  that illustrates its seasonal component. Aggregated demands usually follow a period of one week because human activities are conditioned to the hour of the day and the day of week. Other slow periods, such as annual seasons or tropical temperature oscillations, that depend on the geographic location of the loads, can also be identified following a similar analysis. However, for the sake of simplicity and generality, only a one-weekly period will be considered in this paper, keeping in mind that all subsequent procedures for model identification can be extended to time series with multiple periods following analogous methods.

Fig. 5 shows the estimated autocorrelation function of time series  $\Delta_s z_t = z_t - z_{t-s}$ , with  $s = 2,016$ . A small version of the resulting time series is also shown in Fig. 5. Note that, after removing the periodic component of the series, most of the estimated autocorrelations are reduced. However, strong oscillating autocorrelations remain, suggesting non-stationarity.

The next step in model identification is checking whether the time series is homogeneous. Homogeneity is perceived when different, sufficiently large, intervals of the time series show similar patterns in terms of level and slope. Large portions of  $\Delta_s z_t$  show similar mean levels which indicates a first degree of homogeneity, i.e., each outcome is highly influenced by its previous realization. Thus, Fig. 6 shows the estimated autocorrelation function of the time series  $\Delta \Delta_s z_t = z_t - z_{t-s} - (z_{t-1} - z_{t-s-1})$  and a small version of the resulting time series  $\omega_t = \Delta \Delta_s z_t$ , with  $s = 2,016$ .

Time series  $\omega_t$  in Fig. 6 is a stationary time series whose estimated autocorrelations can be considered significant for lags 0, 1, 2,015, 2,016, and 2,017. Furthermore, lags 1 and 2,016 are negative, whereas lags 2,015 and 2,017 are positive which is consistent with a moving average process given by



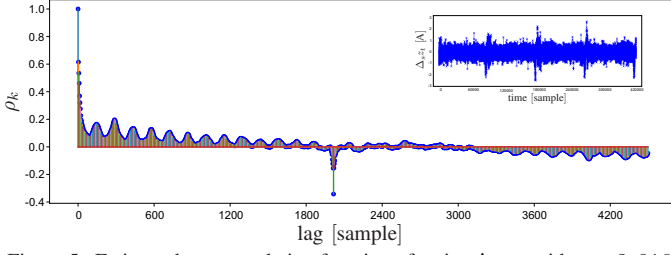


Figure 5. Estimated autocorrelation function of series  $\Delta_s z_t$ , with  $s = 2,016$ .

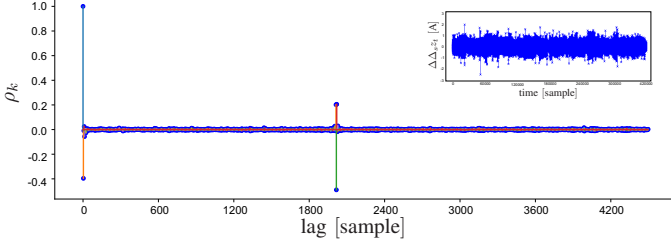


Figure 6. Estimated autocorrelation function of stationary series  $\omega_t = \Delta\Delta_s z_t$ , with  $s = 2,016$ .

the following regression model:

$$\omega_t = \Delta\Delta_s z_t = a_t - \theta a_{t-1} - \Theta a_{t-s} + \theta\Theta a_{t-s-1}, \quad (1)$$

where stochastic variables  $\omega_t$  and  $a_t$ 's have zero mean. Equation (1) is known as univariate multiplicative ARIMA model [20], and only three parameters need to be estimated in order to start generating forecasts with (1): parameters  $\theta$ ,  $\Theta$ , and the variance of the  $a_t$ 's, i.e.,  $\sigma_a^2$ .

### B. Preliminary Estimation

Parameters  $a_t$ 's in (1) are independent and identically distributed (IID) random variables whose linear regression predicts the values of the future outcomes of  $\omega_t$ . Thus, the autocovariance function of (1), for lags  $k = 0, 1, \dots, \infty$ , is given by (2).

$$\gamma_k = E[(a_t - \theta a_{t-1} - \Theta a_{t-s} + \theta\Theta a_{t-s-1})(a_{t-k} - \theta a_{t-k-1} - \Theta a_{t-k-s} + \theta\Theta a_{t-k-s-1})] \quad (2)$$

Since  $a_t$ 's are IID, the analytical autocovariances of  $\omega_t$  are given by (3)–(8). Note that, since  $\rho_k = \frac{\gamma_k}{\gamma_0}$ , (3)–(8) are consistent with the estimated autocorrelations in Fig. 6.

$$\gamma_0 = [1 + \theta^2 + \Theta^2 + (\theta\Theta)^2] \sigma_a^2 = (1 + \theta^2)(1 + \Theta^2) \sigma_a^2 \quad (3)$$

$$\gamma_1 = [-\theta - \Theta(\theta\Theta)] \sigma_a^2 = -\theta(1 + \Theta^2) \sigma_a^2 \quad (4)$$

$$\gamma_{s-1} = \theta\Theta \sigma_a^2 \quad (5)$$

$$\gamma_s = [-\theta - \Theta(\theta\Theta)] \sigma_a^2 = -\theta(1 + \theta^2) \sigma_a^2 \quad (6)$$

$$\gamma_{s+1} = \theta\Theta \sigma_a^2 \quad (7)$$

$$\gamma_k = 0; \text{ otherwise} \quad (8)$$

Providing that the series  $\omega_t$  is stationary, and given the autocovariances in (3)–(8), the preliminary estimation of parameters  $\theta$  and  $\Theta$  can be obtained by (9) and (10), respectively.

$$\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{-\theta}{1 + \theta^2} \quad (9)$$

$$\rho_s = \frac{\gamma_s}{\gamma_0} = \frac{-\theta}{1 + \Theta^2} \quad (10)$$

Thus, based on the empirical results from Fig. 6, the preliminary parameters are  $\hat{\theta} = 0.4916$  and  $\hat{\Theta} = 0.8122$ , where the accent  $\hat{a}$  stands for “estimated value of  $a$ ”. Finally, an unbiased estimation of  $\sigma_a^2$  may be obtained from the time series data and the ARIMA model as discussed in the following subsection.

### C. Forecasting

Once initial values for  $\hat{\theta}$  and  $\hat{\Theta}$  have been obtained, the ARIMA model in (1) can be used to dynamically generate forecasts from the time series. Consider a lag time  $l$ , such that  $l$  can go from 5 minutes-ahead to any practical value of  $l$ . Forecast values  $\tilde{z}_t$  can be produced by sequentially solving (11) for  $t \in \{\tau + 1, \tau + 2, \dots, \tau + l\}$ , where  $\tau$  is the current time from which forecasts are taken and the accent  $\tilde{a}$  stands for “forecasted value of  $a$ ”.

$$\tilde{z}_t = z_{t-1} + z_{t-s} - z_{t-s-1} + a_t - \hat{\theta}a_{t-1} - \hat{\Theta}a_{t-s} + \hat{\theta}\hat{\Theta}a_{t-s-1} \quad \forall t \in \{\tau + 1, \tau + 2, \dots, \tau + l\} \quad (11)$$

In (11), the values for  $z_{t \leq \tau}$  are obtained directly from the series data, whereas the values for  $z_{t > \tau}$  are returned from previous forecasts. On the other hand, the values for  $a_{t \leq \tau}$  are obtained by using the conditional estimation method shown in Algorithm 1, whereas the values for  $a_{t > \tau}$  are set to zero.

#### Algorithm 1 Conditional Estimation of $a_t$ 's

---

```

1:  $e_t \leftarrow 0$ ;  $t \in \{-s-1, -s, \dots, \tau+s, \tau+s+1\}$ 
2:  $\tilde{\omega}_t \leftarrow 0$ ;  $t \in \{-s-1, -s, \dots, -1, 0\}$ 
3:  $a_t \leftarrow 0$ ;  $t \in \{-2s-2, -2s-1, \dots, \tau-1, \tau\}$ 
4: for each  $t \in \{\tau, \tau-1, \dots, 2, 1\}$  do
5:    $e_t = \omega_t + \hat{\theta}e_{t+1} + \hat{\Theta}e_{t+s} - \hat{\theta}\hat{\Theta}e_{t+s+1}$ 
6: for each  $t \in \{-s-1, -s, \dots, -1, 0\}$  do
7:    $\tilde{\omega}_t = e_t - \hat{\theta}e_{t+1} - \hat{\Theta}e_{t+s} + \hat{\theta}\hat{\Theta}e_{t+s+1}$ 
8: for each  $t \in \{-s-1, -s, \dots, \tau-1, \tau\}$  do
9:   if  $t \leq 0$  then
10:     $a_t = \tilde{\omega}_t + \hat{\theta}a_{t-1} + \hat{\Theta}a_{t-s} - \hat{\theta}\hat{\Theta}a_{t-s-1}$ 
11:   else if  $t > 0$  then
12:     $a_t = \omega_t + \hat{\theta}a_{t-1} + \hat{\Theta}a_{t-s} - \hat{\theta}\hat{\Theta}a_{t-s-1}$ 

```

---

Algorithm 1 is called *conditional estimation* because it approximates the individual outcomes of  $a_t$ 's based on the ARIMA model and the historical information of the time series  $\omega_t$ . Considering that the known values of  $\omega_t$  start at  $t = 1$  and end at  $t = \tau$ , back-forecasts of  $\omega_t$  (regarded as  $\tilde{\omega}_t$ ) are first calculated using the recursive procedures in lines 4 to 7, where  $e_t$  are the random coefficients of the reversed series  $\tilde{\omega}_t$ . Once the back-forecast values of  $\tilde{\omega}_{-s-1 \leq t \leq 0}$  have been obtained, the recursive forward method in lines 8 to 12 is used to estimate the  $a_t$ 's. Note that the Algorithm 1 estimates the values of  $a_{-s-1 \leq t \leq \tau}$ , i.e., it also provides the random coefficients of the back-forecasts.

Once the values of  $a_{t \leq \tau}$  have been calculated, the unbiased estimator of the variance  $\hat{\sigma}_a^2$  is obtained by averaging over the sum-of-square values of  $a_t$  given by (12), where  $N$  is the number of samples in the time series  $\omega_t$ .

$$\hat{\sigma}_a^2 \approx \frac{1}{N} \sum_{t=1}^{\tau} a_t^2 \quad (12)$$

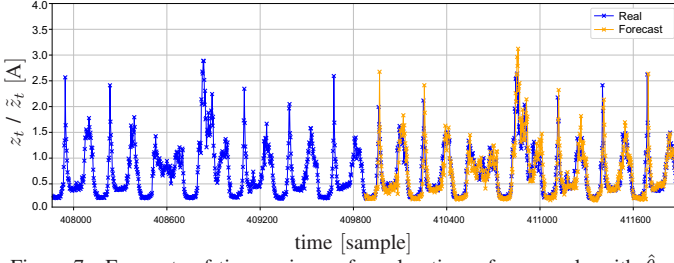


Figure 7. Forecasts of time series  $z_t$  for a lag time of one week, with  $\hat{\theta} = 0.4916$  and  $\hat{\Theta} = 0.8122$ .

Fig. 7 shows the forecasts of the series  $z_t$  in Fig. 2, obtained using the aforementioned technique for a lag  $l = 2,016$ , i.e., one-week ahead. The forecasts in yellow are compared with the real outcomes in blue to demonstrate the accuracy of the proposed ARIMA model when using preliminary values of  $\hat{\theta} = 0.4916$  and  $\hat{\Theta} = 0.8122$ . The estimated variance  $\hat{\sigma}_a^2 \approx 10.2$  is obtained by (12). Note that the forecasted demands follow a similar pattern as the real consumption. Through the estimated variance  $\hat{\sigma}_a^2$ , the standard deviation of the forecast errors at any lag  $l$  can be assessed, and they increase with the length of  $l$  (see Chapter 5 in [20]).

#### IV. MODEL ESTIMATION

Together, the seasonal ARIMA model in (1) and the forecasting algorithm in Section III-C, provide a parsimonious short-term load forecasting technique that can be programmed in each individual meter of an EDS with an AMI. As shown in Fig. 7, rough values of  $\hat{\theta}$  and  $\hat{\Theta}$  already provide suitable forecasts for a one-week lag. However, as more data become available, and considering that consumption is a dynamic stochastic process that continuously reacts to exogenous factors, e.g., new circuits that are installed to energize new demands, socioeconomic activities that change in an area, unexpected events that lead to different operating points and topologies of the EDS, etc. Thus, in order to be parsimonious, the proposed short-term load forecasting method must be able to adapt itself to the ever-changing nature of demands by updating the values of the parameters  $\hat{\theta}$  and  $\hat{\Theta}$  when new measurements of the series are obtained.

Two approaches are presented in this section to optimized the values of  $\hat{\theta}$  and  $\hat{\Theta}$ : a frequentist and a Bayesian approach. These two methods are different strategies to optimize the model's parameters based on historical data. However, each one derives from parallel perspectives of statistics. The frequentist analysis considers parameters as fixed constant values of the estimation model, whereas the Bayesian analysis considers parameters as random variables of the estimator with given prior distributions [28].

##### A. Frequentist Approach

In the frequentist approach, an iterative non-linear programming method is used to optimize the values of  $\hat{\theta}$  and  $\hat{\Theta}$  every time new data is attained, as follows:

**Step 1** Let  $k \leftarrow 0$ . Use initial estimations of  $\hat{\theta}^k \leftarrow \hat{\theta}$  and  $\hat{\Theta}^k \leftarrow \hat{\Theta}$ .

**Step 2** Use Algorithm 1 for calculating the  $a_t^{k*}$ 's of the time series for a set of updated measurements of  $z_t$ , and use (12) to estimate the variance  $\hat{\sigma}_{a^k}^2$ .

**Step 3** Determine the negative derivatives  $x_t^{\hat{\theta}^k}$  and  $x_t^{\hat{\Theta}^k}$ , in which each term is calculated using (13) and (14), for  $t \in \{1, \dots, \tau\}$  and a small  $\delta$ .

$$x_t^{\hat{\theta}^k} = \frac{1}{\delta} \left[ a_t^k |_{\hat{\theta}^k, \hat{\Theta}^k} - a_t^k |_{\hat{\theta}^k + \delta, \hat{\Theta}^k} \right] \quad (13)$$

$$x_t^{\hat{\Theta}^k} = \frac{1}{\delta} \left[ a_t^k |_{\hat{\theta}^k, \hat{\Theta}^k} - a_t^k |_{\hat{\theta}^k, \hat{\Theta}^k + \delta} \right] \quad (14)$$

**Step 4** Update the parameters using the least square estimator as in (15), where coefficients  $f_{11}$ ,  $f_{12}$ ,  $f_{21}$  and  $f_{22}$  are given by (16).

$$\begin{bmatrix} \hat{\theta}^{k+1} \\ \hat{\Theta}^{k+1} \end{bmatrix} = \begin{bmatrix} \hat{\theta}^k \\ \hat{\Theta}^k \end{bmatrix} + \begin{bmatrix} f_{11} \sum_{t=1}^{\tau} x_t^{\hat{\theta}^k} a_t^k + f_{12} \sum_{t=1}^{\tau} x_t^{\hat{\Theta}^k} a_t^k \\ f_{21} \sum_{t=1}^{\tau} x_t^{\hat{\theta}^k} a_t^k + f_{22} \sum_{t=1}^{\tau} x_t^{\hat{\Theta}^k} a_t^k \end{bmatrix} \quad (15)$$

$$\begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \sum_{t=1}^{\tau} (x_t^{\hat{\theta}^k})^2 & \sum_{t=1}^{\tau} x_t^{\hat{\theta}^k} x_t^{\hat{\Theta}^k} \\ \sum_{t=1}^{\tau} x_t^{\hat{\theta}^k} x_t^{\hat{\Theta}^k} & \sum_{t=1}^{\tau} (x_t^{\hat{\Theta}^k})^2 \end{bmatrix} \quad (16)$$

**Step 5** If  $\|\hat{\theta}^{k+1} - \hat{\theta}^k\| < \epsilon$  and  $\|\hat{\Theta}^{k+1} - \hat{\Theta}^k\| < \epsilon$ , then Stop. Otherwise, let  $k \leftarrow k + 1$  and return to Step 2.

The aforementioned frequentist approach was used to optimize parameters  $\hat{\theta}$  and  $\hat{\Theta}$  using the preliminary values obtained in Section III-B as initial guesses, and  $\delta = \epsilon = 0.001$ . After 54 iterations, the optimized values are  $\hat{\theta} = 0.6033$  and  $\hat{\Theta} = 0.9567$ , which has an estimated variance of  $\hat{\sigma}_a^2 = 8.53$ , i.e., 16.4% lower than the same variance obtained with the preliminary values.

##### B. Bayesian Approach

Let  $p(\theta, \Theta)$  be the joint probability distribution function for random variables  $\theta$  and  $\Theta$ , prior to the data. Then, Bayes's theorem in (17) states that the posterior probability distribution function of  $\theta$  and  $\Theta$  given a collection of data  $\mathbf{z}$ , i.e.,  $p(\theta, \Theta | \mathbf{z})$ , is proportional to the product between prior distribution  $p(\theta, \Theta)$  and the joint distribution of the data given parameters  $\theta$  and  $\Theta$ , i.e.,  $p(\mathbf{z} | \theta, \Theta)$ .

$$p(\theta, \Theta | \mathbf{z}) \propto p(\mathbf{z} | \theta, \Theta) p(\theta, \Theta) \quad (17)$$

Assuming  $a_t$ 's and  $w_t$ 's are normally distributed, it can be demonstrated that the joint distribution function of data given parameters  $\theta$  and  $\Theta$  is (18), where  $S(\theta, \Theta) = \sum_{t=1}^{\tau} a_t^2$  is the conditional sum-of-squares function, and  $f(\theta, \Theta)$  is a non-linear function of  $\theta$  and  $\Theta$  [20].

$$p(\mathbf{z} | \theta, \Theta) = f(\theta, \Theta) \exp \left\{ -\frac{1}{2\hat{\sigma}_a^2} S(\theta, \Theta) \right\} \quad (18)$$

Using Jeffery's prior [29] for both parameters, and assuming a constant value for  $\hat{\sigma}_a^2$ , the posterior probability distribution function  $p(\theta, \Theta | \mathbf{z})$  has the form of (19), where  $N$  is the number of data samples of the time series  $\mathbf{z}$ .

$$p(\theta, \Theta | \mathbf{z}) \propto [S(\theta, \Theta)]^{-N/2} \quad (19)$$

Taking samples from  $0 \leq \theta \leq 1$  and  $0 \leq \Theta \leq 1$ , and using Algorithm 1 for calculating the  $a_t^{k*}$ 's, the posterior

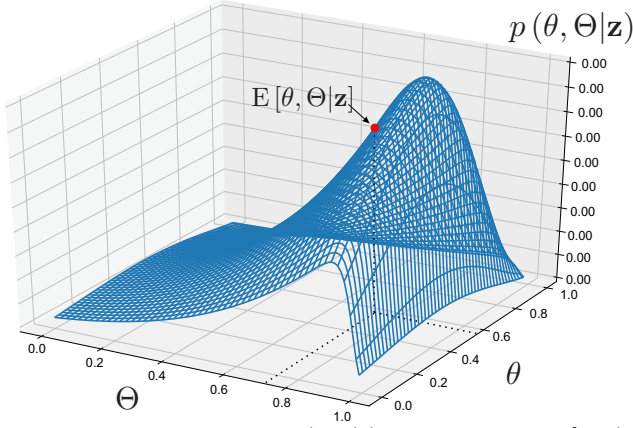


Figure 8. Joint PDF of the posterior  $p(\theta, \Theta|z)$  and expected value  $E[\theta, \Theta|z]$ .

probability distribution function  $p(\theta, \Theta|z)$  in (19) can be plotted to identify the mean values and statistical moments of  $\theta$  and  $\Theta$ , using the most recent available information of  $z$ . Fig. 8 shows the resulting joint probability distribution of  $p(\theta, \Theta|z)$  which not only indicates the expected values of  $\hat{\theta} = 0.5421$  and  $\hat{\Theta} = 0.7274$  but also, it provides a confidence interval of 95% for  $0.4 \leq \theta \leq 0.8$  and  $0.7 \leq \Theta \leq 1.0$ .

### C. Adaptability

In practice, the process of updating the values of  $\hat{\theta}$  and  $\hat{\Theta}$  using either the frequentist or Bayesian approach is done independently for each time series (and smart meter) as follows: whenever a new realization of the time series is available at time  $\tau$ , use the previous  $N$  samples to deploy the model estimation approach of choice and disregard any sample before  $\tau - N$ . The forecasted values of  $\hat{z}_t$  are only calculated when the OOP requires them using (11) for a lag time  $l$ .

## V. MODEL ADEQUACY

Low values of the estimated variances  $\hat{\sigma}_a^2$  are not a sufficient indication of the forecasting method's accuracy. Checking the autocorrelation function of the residuals is also an important aspect because high autocorrelations indicate that critical information from the original series has been left behind by the model. Figs. 9 and 10 show the estimated autocorrelation function of the residuals  $a_t$ 's for both model estimation approaches, frequentist and Bayesian, considering lags between 1 to 4,500 samples (i.e., up to two weeks). Despite some cyclical deviations, which are expected in random series, there are no noticeable large autocorrelations that indicate an evident lack-of-fit from both methods.

A more systematic way for checking the adequacy of the model is to assess the statistical significance of apparent

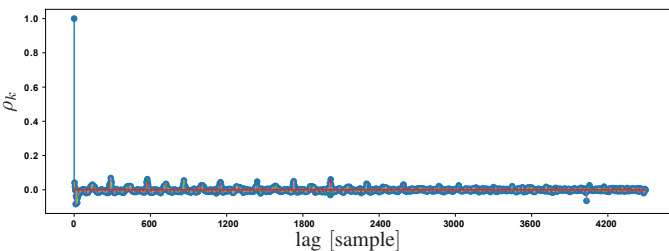


Figure 9. Frequentist approach: Estimated autocorrelation function of the residuals  $a_t$ 's.

deviations of the residuals. To do so, the Ljung-Box-Pierce [30] statistic test can be performed using (20).

$$\bar{Q} = N(N+2) \sum_{k=1}^K (N-k)^{-1} \rho_k^2 \quad (20)$$

where  $N$  is the interval of the data series used to fit the model and  $K$  is the interval of the lags under study. If the model is appropriate the statistic  $\bar{Q}$  is approximately distributed as  $\chi^2(K-p-q)$ , where  $p = 0$  and  $q = 2$  in the ARIMA model (1). Thus, considering  $K = 4,032$  (i.e., two weeks) and fitting the model using  $N = 34,944$  samples (i.e., one year) the value of  $\bar{Q}$  was approximately 3,032 using the frequentist approach, and 3,472 for the Bayesian approach. Both  $\bar{Q}$ 's are below the adequacy level of 4,179 that corresponds to the 5%- $\chi^2$  test, with 4,030 degrees of freedom. In his case, the Ljung-Box-Pierce statistic test does not provide any evidence of inadequacy in the model.

## VI. TESTS AND RESULTS

In order to test the efficiency of the proposed short-term load forecasting technique in the context of solving OOP problems, a robust self-healing scheme was simulated. To that end, the real-size 13.2 kV EDS in Fig. 11 was used for tests. The system comprises five radial feeders that supply electricity to 38,000 users of three types: residential, commercial and industrial consumers. Individual load profiles were randomly generated using the LoadProfileGenerator software in [26], [27]. The real-size EDS has 5,181 nodes, from which 955 are primary distribution transformers with smart meters. The meters are constantly measuring average three-phase current and voltage magnitudes with a resolution of 5 minutes per sample. Thus, each smart meter supervises at least six time series per transformer and, it is assumed that at least four years of measurements are available to the OOP system. Blue and magenta feeders in Fig. 11 are connected to a main substation of 40 MVA (nominal capacity), whereas red, yellow, and green feeders are connected to a different substation of 60 MVA. Moreover, the EDS has three 5 MVA dispatchable distributed generation (DG) units with islanded operation capabilities. The total installed capacity of the distribution transformers is approximately 81 MVA. Also, there are 32 remotely-controlled normally closed (NC) switches and 14 remotely-controlled normally open (NO) switches that participate in the self-healing scheme.

The proposed short-term load forecasting model in (1) is used to predict the three-phase consumption at each transformer. Thus, all the 955 meters in Fig. 11 are set to be constantly measuring at least six different time series (three-phase voltage magnitudes and load currents) and deploying

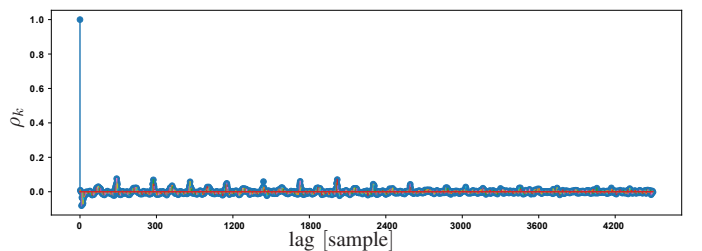


Figure 10. Bayesian approach: Estimated autocorrelation function of the residuals  $a_t$ 's.

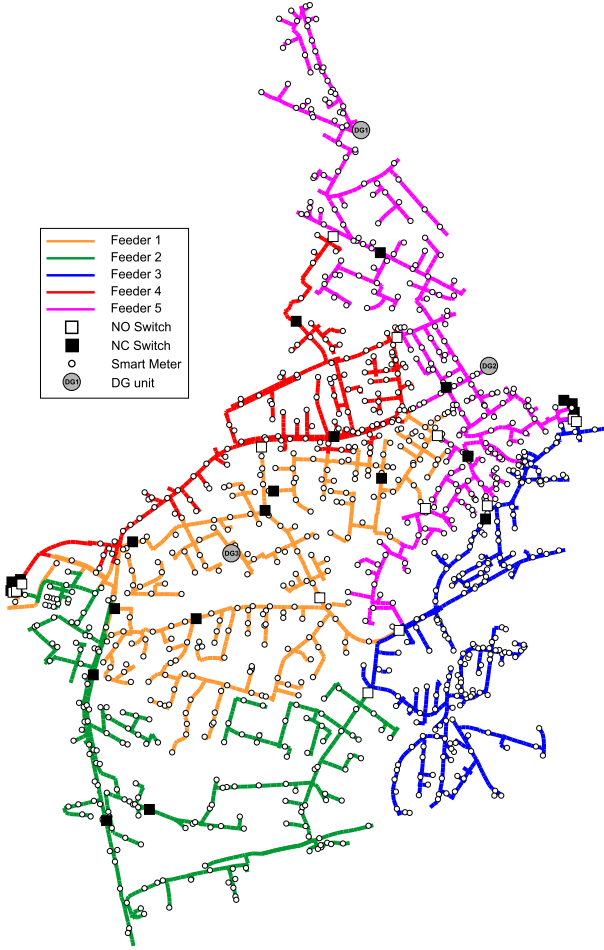


Figure 11. Real-size 13.2kV EDS with an AMI.

one-week ahead forecasts with a lag of  $l = 2,016$  samples, whenever the self-healing scheme requires them. Note that, in theory, any lag  $l$  can be used, but the standard deviation of the forecast errors increases with the length of  $l$  [20]. Hence, the proposed ARIMA model can only be employed for short-term planning, because the forecasted data for the next week will have larger errors than the forecasts for the next day.

Two fault scenarios are discussed in sections VI-A and VI-B, respectively. In Case 1, the restoration after a permanent fault of the main breaker at feeder 5 (magenta) is simulated. In Case 2, a permanent fault of the main breaker of feeder 1 (yellow) is deployed. In both cases, the fault requires one day to be totally repaired. Thus, only the first 288 forecasts are used to deploy the self-healing system.

For each case, the two proposed methods used to estimate parameters  $\theta$  and  $\Theta$ , i.e., the frequentist approach in Section IV-A and the Bayesian approach in Section IV-B, are compared with a basic restoration plan obtained either by using the nominal capacities of all transformers as conventional demands, a.k.a., the *nominal approach*, or the load diagram of the previous day as an estimation of future consumption, a.k.a., the *day-before approach*.

Finally, as mentioned in Section II-A, the centralized self-healing scheme was deployed using an heuristic solution of the restoration model presented in [25]. The heuristic is an adaptation of the Tabu Search algorithm [31] that returns the

best restoration sequence found after 60 seconds. In all cases, the objective function minimizes the unsupplied demand and the number of switch operations, and penalizes the violations of the operational limits. The Tabu Search attribute has been set to 1 for all simulations (see [31] for more information).

#### A. Case 1: Fault at the main breaker of feeder 5

A permanent fault of the main breaker in feeder 5 (magenta) disconnects all loads downstream the circuit. Thus, a restoration sequence is deployed using the load forecasts from four different approaches, as follows:

1) *Nominal approach*: In this case, demands are considered as constant balanced three-phase active and reactive power injections, equal to the nominal capacities of the distribution transformers with an inductive power factor of 0.9. Thus, given a permanent fault of the main breaker of feeder 5 (magenta), the solution generated by the self-healing scheme using nominal capacities is shown in Fig. 12a. Note that, using the nominal capacities of the transformers, the self-healing scheme was not able to find a feasible solution that minimizes the unsupplied demand. Thus, as shown in Fig. 12a, all nodes and circuits of feeder 5 were de-energized by the opening of the main breaker at the substation, and no further restoration actions were performed.

2) *Day-before approach*: As an alternative to the nominal approach, a more simplistic method based on the load profile curve of the previous day can be used to predict the moment of maximum consumption after the fault. The solution generated by the self-healing scheme using the day-before approach is shown in Fig. 12b. In this case, the moment of maximum consumption of the previous day was 8% lower than the actual realization of the moment of maximum consumption after the fault. Thus, even though the solution in Fig. 12b transferred all demands from feeder 5 to feeder 3 (blue), the underestimation of the day-before approach produced several lines with overcurrent limits transgressions, mostly concentrated at the beginning of feeder 3, for approximately one hour during the post-fault operation. The mean absolute percentage error (MAPE) of this approach for all the 955 meters considering a one-day ahead forecast was 16%.

3) *Frequentist approach*: Clearly, assuming nominal capacities is very conservative because it is unrealistic to believe that all demands require the complete installed capacity of the EDS, simultaneously. Thus, after using the frequentist approach to generate forecasts, the moment of maximum consumption was identified by comparing the maximum forecasted currents and the minimum forecasted voltages in Fig. 13 and Fig. 14. Note that, since the expected time to repair the outage is one day, only the blue area of Fig. 13 and Fig. 14 were considered. In this case, the moment of maximum consumption occurs at lag  $l = 78$ , i.e., approximately after 6.5 hours after the outage. Thus, using the demands from the moment of maximum consumption, plus the standard deviation of the forecast errors provided by the ARIMA model, the restoration plan shown in Fig. 12d was obtained by the centralized self-healing scheme. In this case, all demands were supplied after transferring most of the loads from feeder 5 to feeder 3 (blue), and the remaining loads from feeder 5



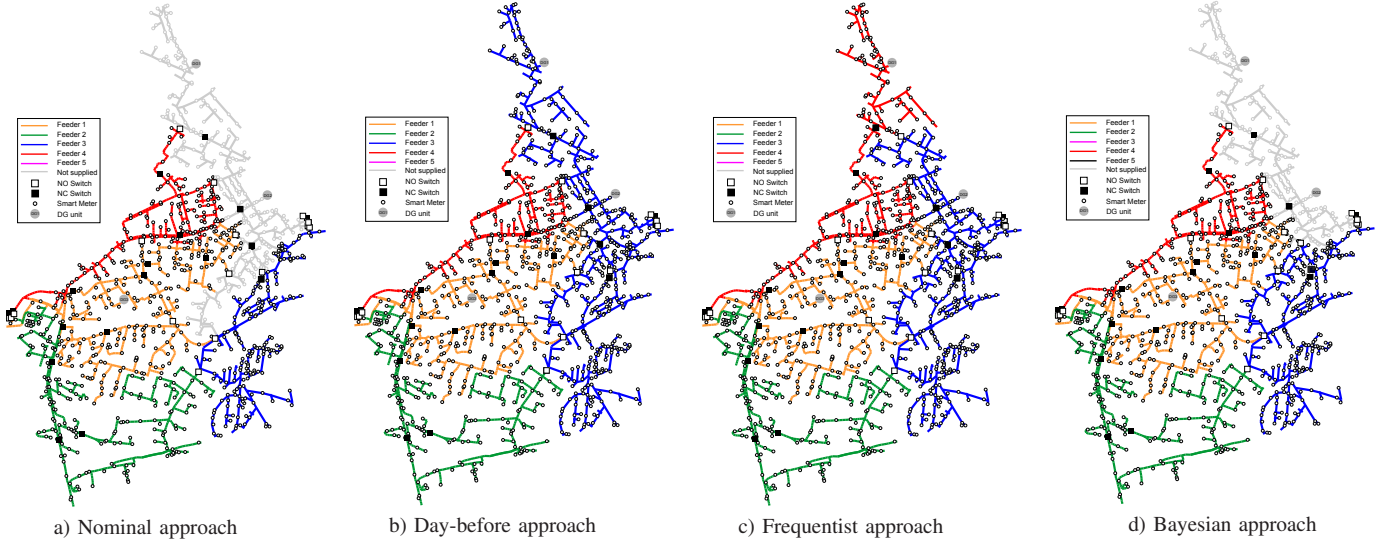


Figure 12. Self-healing solution for Case 1: a) using nominal approach; b) day-before approach; c) frequentist approach; d) Bayesian approach.

to feeder 4 (red). During post-fault operation, three lines of feeder 3 have slight overcurrent of 109% of their maximum current capacities. The one-day ahead MAPE of this approach was 12%.

4) *Bayesian approach*: In this case, the Bayesian approach is used to generate the forecasts, and the moment of maximum consumption is identified as before. With the Bayesian approach, the moment of maximum consumption occurs at lag  $l = 81$ , i.e., approximately after 6.75 hours after the outage. Thus, using the demands from the moment of maximum forecasted consumption, plus the standard deviation of the forecast errors, the solution shown in Fig. 12c was obtained by the self-healing scheme. Unlike the frequentist approach, the restoration plan using the Bayesian estimation did not supplied all demands from feeder 5. Instead, only a portion of the consumers were transferred to feeder 3 (blue), while the rest remained de-energized. However, in terms of number of de-energized consumers, the Bayesian solution was still better than the solution using nominal transformer capacities, with no limits transgressions during post-fault operation. The one-day ahead MAPE of this approach was 12%.

A summary of all four restoration plans in Fig. 12 is shown in Table I. The minimum voltage, maximum current and limits transgressions are taken from the post-fault operation. Thus, as evidenced by the results in Table I, the impact and efficiency of the self-healing scheme is highly improved when suitable forecasted demands are used instead of nominal references or simplistic previous days reproductions.

Table I  
SELF-HEALING RESULTS FOR EACH APPROACH IN FIG. 12.

Approach	Nominal	Day-before	Frequentist	Bayesian
Unsupplied demand [MW]	30.56	0.0	0.0	20.62
De-energized users [#]	9441	0	0	7505
Switch operations [#]	0	1	3	2
Minimum voltage [p.u.]	0.9802	0.9706	0.9788	0.9802
Maximum current [A]	417	502	458	417
Limits transgression	None	Eighteen lines with overcurrent	Three lines with overcurrent	None

#### B. Case 2: Fault at the main breaker of feeder 1

A permanent fault of the main breaker in feeder 1 (yellow) disconnects all loads downstream the circuit. The restoration plans for the four different load forecasting methods are shown in Fig. 15. Note that, using the nominal approach, the solution provided by the service restoration method has created an island to supply a fraction of the users in feeder 1. As shown in Fig. 15a, the island used DG unit 3 as the main supply, and it was isolated due to the opening of two NC switches. On the other hand, in the day-before, frequentist and Bayesian approaches, all users in feeder 1 were transferred to different

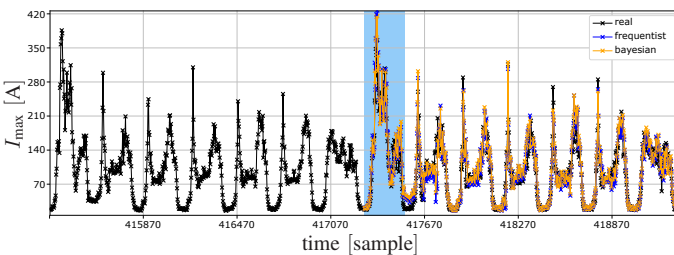


Figure 13. Realization (black), frequentist (blue) and Bayesian (yellow) forecasts of the maximum current magnitude.

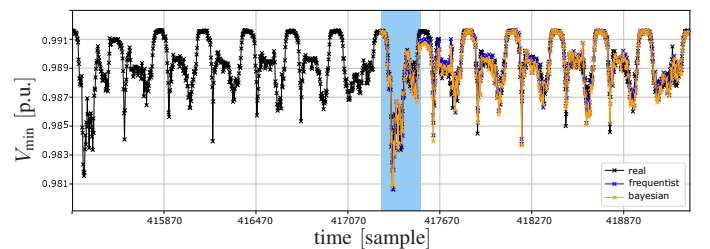


Figure 14. Realization (black), frequentist (blue) and Bayesian (yellow) forecasts of the minimum voltage magnitude.

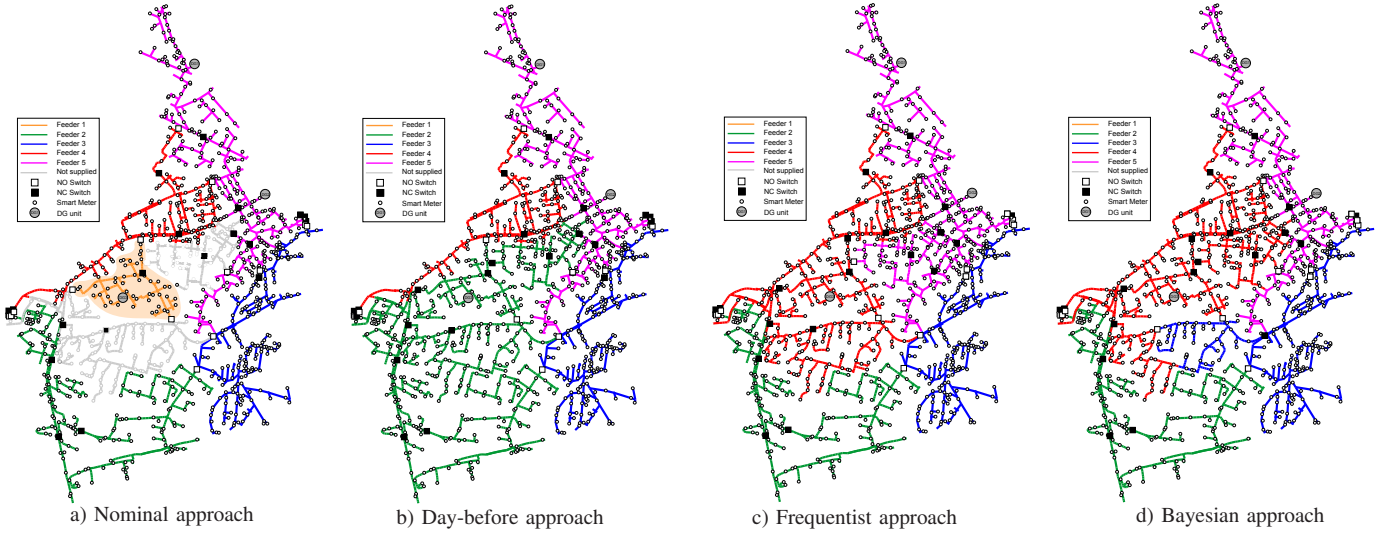


Figure 15. Self-healing solution for Case 2: a) using nominal approach; b) day-before approach; c) frequentist approach; d) Bayesian approach.

Table II  
SELF-HEALING RESULTS FOR EACH APPROACH IN FIG. 15.

Approach	Nominal	Day-before	Frequentist	Bayesian
Unsupplied demand [MW]	15.23	0.0	0.0	0.0
De-energized users [#]	10697	0	0	0
Switch operations [#]	2	1	3	3
Minimum voltage [p.u.]	0.9802	0.9680	0.9709	0.9744
Maximum current [A]	417	455	430	417
Limits transgressions	None	Three lines with overcurrent	None	None

feeders, but with unique configurations for each case. As shown in Table II, the underestimation of the forecasted loads provided by the day-before approach led to overcurrent events during the post-fault operation in feeder 2. Finally, the solutions obtained with the frequentist and the Bayesian approaches supplied all demands of feeder 1 and they did not lead to any operational limits transgressions during the post-fault operation.

### C. Frequentist approach vs Bayesian approach

In this case, if future outcomes of the series  $\omega_t$  behave mostly as a Gaussian process with zero mean and lower values of  $\hat{\sigma}_a^2$ , then the frequentist approach should be used because it is based on the least-square estimator that minimizes the variance of the forecast errors. However, if large portions of the series  $\omega_t$  present non-homogeneous behavior (see Section III-A) or large values of  $\hat{\sigma}_a^2$  are obtained, then the non-informative nature of Jeffery's prior used in the Bayesian approach could produce more accurate results because there might be an underlying randomness in the nature of  $\Theta$  and  $\theta$ , that is not captured by the frequentist approach. Thus, a good rule-of-thumb would be to use the frequentist approach first, and then, depending on the range of the recurring values of

$\hat{\Theta}$  and  $\hat{\theta}$ , the Bayesian approach can be deployed using tight intervals for both unknown parameters.

From the statistical point-of-view, there are no good reasons for using one inference method over the other [28]. However, some practical aspects might be considered for implementation. For example, the convergence of the iterative process in Section IV-A is not always guaranteed, whereas the Bayesian estimation does not rely on any convergence process. One major drawback of the Bayesian approach in Section IV-B is that, for large data samples, the posterior in (19) can result in extremely low numbers that cannot be computationally handled, thus lower values of  $N$  must be required.

## VII. CONCLUSIONS

Assuming three-phase electrical measurements as univariate non-stationary seasonal time series, this paper investigates two parsimonious short-term load forecasting techniques to estimate future aggregated demands for the OOP of EDS: a frequentist and a Bayesian multiplicative seasonal ARIMA model. Both approaches are shown to satisfy the four parsimonious requirements: low computational complexity, fair accuracy in terms of expected value and residuals, few tuning parameters, and adaptability. Thus, they are suitable for solving OOP problems in highly supervised EDS. Results show that the use of the proposed parsimonious short-term forecasting techniques have a significant impact on the quality of the OOP methods, such as the centralized self-healing scheme, especially compared with other approaches that use nominal references or simplistic day-before approaches.

## APPENDIX

The optimal restoration switching sequence for unbalanced three-phase EDS is given by the mixed integer non-linear programming (MINLP) model in (21)–(28).

$$\min \sum_{z \in \Omega_z} \sum_{s \in \Omega_s} c_{z,s}^U (1 - x_{z,s}) + c^{sw} \sum_{km \in \Omega_{sw}} \sum_{s \in \Omega_s} (\Delta y_{km,s}^+ + \Delta y_{km,s}^-) \quad (21)$$

Subject to:

$$\mathbf{F} \left( I_{ij,f,s}^{r,i}, I_{km,f,s}^{sw,r,i}, V_{i,f,s}^{r,i}, P_{i,f}^D, Q_{i,f}^D, x_{z,s}, y_{km,s}, P_{g,s}^{DG}, Q_{g,s}^{DG} \right) = 0$$

$$\forall i \in \Omega_b, ij \in \Omega_l, km \in \Omega_{sw}, f \in \Omega_f, g \in \Omega_g, s \in \Omega_s, z \in \Omega_z \quad (22)$$

$$\mathbf{G} \left( I_{ij,f,s}^{r,i}, I_{km,f,s}^{sw,r,i}, V_{i,f,s}^{r,i}, P_{i,f}^D, Q_{i,f}^D, x_{z,s}, y_{km,s}, P_{g,s}^{DG}, Q_{g,s}^{DG} \right) \leq 0$$

$$\forall i \in \Omega_b, ij \in \Omega_l, km \in \Omega_{sw}, f \in \Omega_f, g \in \Omega_g, s \in \Omega_s, z \in \Omega_z \quad (23)$$

$$y_{km,s} - y_{km,s-1} = \Delta y_{km,s}^+ - \Delta y_{km,s}^-$$

$$\forall km \in \Omega_{sw}, s \in \Omega_s | y_{km,0} = s_{km}^{ini} \quad (24)$$

$$\sum_{km \in \Omega_{sw}} \left( \Delta y_{km,s}^+ + \Delta y_{km,s}^- \right) \leq 1 \quad \forall s \in \Omega_s \quad (25)$$

$$|x_{\hat{z}_k,s} - x_{\hat{z}_m,s}| \leq 1 - y_{km,s} \quad \forall km \in \Omega_{sw}, s \in \Omega_s \quad (26)$$

$$\sum_{km \in \Omega_{sw}} y_{km,s} x_{\hat{z}_k,s}^{\max} = \sum_{z \in (\Omega_z \setminus \Omega_z^S)} x_{z,s}^{\max} \quad (27)$$

$$y_{km,s}, x_{z,s}, \Delta y_{km,s}^{+,-} \in \{0, 1\}$$

$$\forall km \in \Omega_{sw}, z \in \Omega_z, s \in \Omega_s \quad (28)$$

For the sake of simplicity, details regarding the definition of sets, variables, parameters and DG units' operation are omitted here, and can be consulted in [25]. The objective function in (21) minimizes the unsupplied demand using a cost of de-energization  $c_{z,s}^U$ , plus the number of switching operations using a cost of switching  $c^{sw}$ . Unbalanced three-phase power flow equations and operating limits are contained in (22) and (23). Note that the binary decision variables  $x_{z,s}$  and  $y_{km,s}$  influence the outcomes of both functions, where  $x_{z,s}$  represents the status of a given zone  $z$  at restoration step  $s$  (i.e., either energized or de-energized), and  $y_{km,s}$  represents the status of switch  $km$  at restoration step  $s$  (i.e., either open or closed). Given an initial status of the switches  $s_{km}^{ini}$ , (24)–(26) establishes the relationship between the status of a given zone and the operation of the switches at each step of the restoration process. Particularly, (25) guarantees that only one operation can be deployed in each step of the sequence. Constraint (27) guarantees that the topology at the last step of the process ( $s^{\max}$ ) must be radial. The binary nature of the decision variables  $y_{km,s}$ ,  $x_{z,s}$  and  $\Delta y_{km,s}^{+,-}$  is given by (28). The MINLP model in (21)–(28) is a non-convex combinatorial optimization problem and, for practical applications, quality solutions can be obtained via modern heuristics within reasonable computational times and resources.

#### VIII. ACKNOWLEDGMENTS

This work was supported by the Brazilian institution FAPESP-research grant 2017/02196-0.

#### REFERENCES

- [1] J. W. Taylor and P. E. McSharry, "Univariate methods for short-term load forecasting," in *Advances in Electric Power and Energy Systems: Load and Price Forecasting*, M. E. El-Hawary, Ed. New York: Wiley, 2017, ch. 2, pp. 17–39.
- [2] T. Hong, "Energy forecasting: Past, present, and future," *Foresight: The International Journal of Applied Forecasting*, no. 32, pp. 43–48, 2014.
- [3] J. Northcote-Green and R. Wilson, *Control and Automation of Electrical Power Distribution Systems*. Boca Raton, FL: CRC Press, 2017.

- [4] X. Chen, W. Wu, and B. Zhang, "Robust restoration method for active distribution networks," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4005–4015, Sept. 2016.
- [5] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *Int. J. Forecasting*, vol. 32, no. 3, pp. 914–938, Sept. 2016.
- [6] L. Hernandez et al., "A survey on electric power demand forecasting: Future trends in smart grids, microgrids and smart buildings," *IEEE Commun. Surveys Tutorials*, vol. 16, no. 3, pp. 1460–1495, Mar. 2014.
- [7] A. D. Papalexopoulos and T. C. Hesterberg, "A regression-based approach to short-term system load forecasting," *IEEE Trans. Power Systems*, vol. 5, no. 4, pp. 1535–1547, Nov. 1990.
- [8] S. Fan and R. J. Hyndman, "Short-term load forecasting based on a semi-parametric additive model," *IEEE Trans. Power Systems*, vol. 27, no. 1, pp. 134–141, Feb. 2012.
- [9] Y. Goude, R. Nedellec, and N. Kong, "Local short and middle term electricity load forecasting with semi-parametric additive models," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 440–446, Jan. 2014.
- [10] J. W. Taylor and P. E. McSharry, "Short-term load forecasting methods: An evaluation based on european data," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 2213–2219, Nov. 2007.
- [11] M. T. Hagan and S. M. Behr, "The time series approach to short term load forecasting," *IEEE Trans. Power Syst.*, vol. 2, no. 3, pp. 785–791, Aug. 1987.
- [12] R. Weron, *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. New York: Wiley, 2013.
- [13] H. S. Hippert, C. E. Pedreira, and R. C. Souza, "Neural networks for short-term load forecasting: a review and evaluation," *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 44–55, Feb. 2001.
- [14] A. Khotanzad, R. Afkhami-Rohani, and D. Maratukulam, "Annstf-artificial neural network short-term load forecaster generation three," *IEEE Trans. Power Syst.*, vol. 13, no. 4, pp. 1413–1422, Nov. 1998.
- [15] K.-B. Song, Y.-S. Baek, D. H. Hong, and G. Jang, "Short-term load forecasting for the holidays using fuzzy linear regression method," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 96–101, Feb. 2005.
- [16] T. Hong and P. Wang, "Fuzzy interaction regression for short term load forecasting," *Fuzzy Optimization and Decision Making*, vol. 13, no. 1, pp. 91–103, Mar. 2014.
- [17] B.-J. Chen, M.-W. Chang, and C.-J. Lin, "Load forecasting using support vector machines: a study on eunite competition 2001," *IEEE Trans. Power Syst.*, vol. 19, no. 4, pp. 1821–1830, Nov. 2004.
- [18] S. B. Taieb and R. J. Hyndman, "A gradient boosting approach to the kaggle load forecasting competition," *International Journal of Forecasting*, vol. 30, no. 2, pp. 382–394, June 2014.
- [19] J. R. Lloyd, "Gefcom2012 hierarchical load forecasting: Gradient boosting machines and gaussian processes," *International Journal of Forecasting*, vol. 30, no. 2, pp. 369–374, June 2014.
- [20] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis*, 4th ed. New York: Wiley, 2008.
- [21] T. Hong, P. Pinson, and S. Fan, "Global energy forecasting competition 2012," *Int. J. Forecasting*, vol. 30, no. 2, pp. 357–363, 2014.
- [22] R. R. Mohassel, A. Fung, F. Mohammadi, and K. Raahemifar, "A survey on advanced metering infrastructure," *International Journal of Electrical Power & Energy Systems*, vol. 63, pp. 473–484, Dec. 2014.
- [23] P. L. Cavalcante et al., "Centralized self-healing scheme for electrical distribution systems," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 145–155, Jan. 2016.
- [24] J. R. Aguero, "Applying self-healing schemes to modern power distribution systems," in *2012 IEEE Power and Energy Society General Meeting*, San Diego, CA, July 2012, pp. 1–4.
- [25] J. C. Lopez, J. F. Franco, M. J. Rider, and R. Romero, "Optimal restoration/maintenance switching sequence of unbalanced three-phase distribution systems," *IEEE Trans. Smart Grid*, 2017, to be published.
- [26] N. Pflugrad and U. Muntwyler, "Synthesizing residential load profiles using behavior simulation," *Energy Procedia*, vol. 122, pp. 655–660, 2017.
- [27] N. Pflugrad, J. Teuscher, B. Platzer, and W. Schufft, "Analysing low-voltage grids using a behaviour based load profile generator," in *International Conference on Renewable Energies and Power Quality*, Bilbao, Spain, March 2013, pp. 1–5.
- [28] J. Wakefield, *Bayesian and Frequentist Regression Methods*. New York: Springer, 2013.
- [29] H. Jeffreys, *Theory of Probability*, 3rd ed. Clarendon: Oxford, 1961.
- [30] G. M. Ljung and G. E. P. Box, "On a measure of lack of fit in time series models," *Biometrika*, vol. 65, no. 2, pp. 297–303, Aug. 1978.
- [31] F. Grover and M. Laguna, *Tabu Search*, 1st ed. Boston, MA, USA: Springer, 1997.